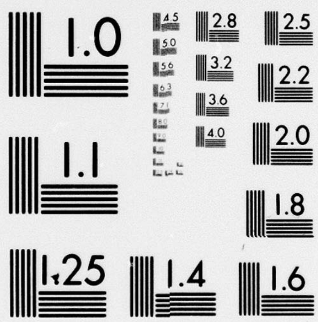
AD-A079 175 ARMY PERSONNEL RESEARCH OFFICE WASHINGTON DC F/G 14/2
FEASIBILITY OF A PROGRAMMED TESTING MACHINE,(U)
NOV 64 A G BAYROFF
UNCLASSIFIED APRO-RESEARCH STUDY-64-3 NL

1 OF 1
AD A
079175

END
DATE
FILMED
2-80
DDC



MICROCOPY RESOLUTION TEST CHART

ADA079175

LEVEL II

Research Study 64-3

FEASIBILITY OF A PROGRAMMED TESTING MACHINE

APRO

U. S. ARMY PERSONNEL RESEARCH OFFICE

Army Project Number
2J024701A722

Input Quality-g

14 APRO-Research Study-64-3

6 # FEASIBILITY OF A PROGRAMMED TESTING MACHINE

10 A. G. Bayroff

Submitted by

Edmund F. Fuchs
Chief, Military Selection Research Laboratory

Accession For	
NTIS GRA&I	☑
DDC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist A	Avail and/or special

Approved by

Charles S. Gersoni
Colonel, GS,
Commanding

J. E. Uhlaner
Director,
Research Laboratories

12 26

Research Studies are special reports to military management. They are usually prepared to meet requests for research results bearing on specific management problems. A limited distribution is made--primarily to the operating agencies directly involved.

11 Nov[redacted]1964

040 650

JOB

FEASIBILITY OF A PROGRAMMED TESTING MACHINE

BRIEF

Requirement:

To investigate the feasibility of applying computer techniques to personnel measurement research and psychometric testing.

Procedure:

APRO research scientists developed a set of requirements which programmed testing equipment would have to meet in order to be useful in research with both conventional tests and a variety of novel test techniques including self-programming or "branching" tests. Working from these requirements, the National Bureau of Standards completed a preliminary design study to determine whether a testing machine meeting the requirements could be constructed.

Findings:

Design of a computerized testing machine suitable for research use has been developed. A less elaborate form of the same design would be suitable for a prototype machine for experimentation under operational conditions. Application of the machine to "branching" tests was worked out in detail.

Utilization of Findings:

Applied in research, a machine constructed according to the design would (1) open new test techniques such as self-programming tests for study and (2) generate test information not now obtainable except through individual administration--for example, time spent on each item, or number of tentative responses as opposed to final answer. Success in further research might well lead to operational testing machines which would improve testing while reducing personnel requirements for test administration.

FEASIBILITY OF A PROGRAMMED TESTING MACHINE

The present report describes a study of the feasibility of applying automation and computer techniques to psychometric testing. Successful application of machine testing techniques could open up new avenues to testing research and also holds the possibility of facilitating testing operations.

As a psychometric tool, the conventional printed test has been invaluable in the measurement of cognitive aptitudes. Only through paper-and-pencil instruments has the Army been able to administer tests to large groups in a uniform and economical manner. Attempts to increase the predictive efficiency of tests and to extend their measurement to aspects of individual behavior not now accessible through large-scale testing have prompted the exploration of programmed test techniques. Such techniques could aid in overcoming some of the constraints imposed by the conventional printed test.

Constraints are, of course, necessary if test scores are to have a uniform interpretation. The single program for all examinees--the same conditions of administration, the same instructions for operating on the items, the same items in the same order--provides such constraints. However, these constraints may exact a price. They may be wasteful in that examinees of high ability must answer items which are too easy for them. Conversely, low ability examinees must attempt items which are too difficult for them. Error may be introduced through carelessness of the able in answering the easy items and through the chance success of the less able in answering the very difficult items. Finally--and of most importance to the research psychologist--the constraints of the printed test prevent acquisition of such information as response time, method of choosing an alternative, the motivating effect of knowledge of results. Whether such added information is useful for prediction would, of course, have to be determined.

The restriction and distortion of information as a result of the constraints inherent in the printed aptitude test may be reduced by resorting to individual test administration. Obviously, this cure is worse than the disease, especially if applied to the testing of large numbers of examinees. Hence, the decision to investigate the possibilities of machine testing.

Should machine testing progress to the point of operational use, certain tangible gains in the efficiency of screening and classification procedures could also be expected:

1. Immediate availability of test scores would facilitate scheduling of steps in the processing system.

2. Human error in the scoring of answer sheets would be eliminated.

3. Testing time would be reduced; or additional tests could be administered in the same amount of time.

4. Number of personnel engaging in testing could be reduced.

5. Test administration would be more uniform.

Investigation took the form of a feasibility study (preliminary design study) to determine whether a machine meeting research and operational requirements can be designed and constructed. The study was made by the National Bureau of Standards on contract with the U. S. Army Personnel Research Office. The study resulted in a conclusion that such machine is feasible and provided a conceptual design utilizing available elements. The design is a product of close collaboration between the two agencies.[1]

REQUIREMENTS FOR A PROGRAMMED TESTING MACHINE

Prior to the decision to undertake a study of the feasibility of designing a suitable testing machine, a review was made of available teaching machines and simulators to determine if any existed which could be used, as is or with modification, for testing research and operations. None was found which could be used as is, and all would require extensive modification. The limited applicability of existing equipment to testing, and the costs of the necessary modifications, led to the conclusion that it would be much less expensive to develop a machine specifically for use in testing.

Before a design study of a programmed testing machine could be undertaken, it was necessary to develop a set of requirements to be met by a machine. The general approach in developing the requirements was to consider primarily testing technique and information, with only secondary regard for feasibility, except where it was known that engineering costs and complexity would far outweigh the likely value of the information generated.

For the sake of clarity, separate requirements were developed for two types of machine. One type would present a fixed or linear program test--that is, a conventional test in which all examinees answer the

---

[1] The report submitted by the National Bureau of Standards, "Report on a Design Study of a Programmed Testing Machine," 18 March 1964, Contract ARO-4, is on file in the U. S. Army Personnel Research Office.

same items in the same order--but yielding additional test information not obtainable with the conventional printed test. The second would present a variable program test, a technique not feasible with the conventional printed test, and on which only exploratory studies have been made.[2] The variable program test technique permits changing the test program according to the performance of the examinee--if the response to an item is correct, a more difficult item is presented next; if incorrect, a less difficult item is presented next. This technique (comparable to that used in the branching of programmed teaching machines) permits arrival at threshold levels for all but the lowest ability examinees with fewer items than the conventional power test.

During discussions with National Bureau of Standards personnel, the point of departure was the set of requirements for the variable program machine. As these were met, the requirements for the fixed program machine were added. The preliminary design study considered all requirements for both types of test. A machine based on the resulting design could be used for research with both the conventional linear program test and a variety of branching tests.

For the laboratory machine, requirement was for as flexible a research capability as was feasible in order not to restrict unnecessarily the scope of research that could be undertaken. For an experimental prototype operational machine, where extensive flexibility is not needed, a reduced set of requirements was established. Thus, the machine as designed for possible operational use would be simpler and less expensive than the laboratory model.

CAPABILITIES OF THE MACHINE DESIGNED

In a machine built according to the design provided, any type of multiple-choice question and instructions to the examinee that have appeared in printed test booklets may be used. Items can be verbally or graphically stated. The number of possible choices per item is greater than four and may be preselected individually for each item. The use of more than four alternatives per item would reduce the error due to guessing correctly. A maximum time signal may be used for each item to warn the examinee he is taking too much time.

---

[2] Seeley, Leonard C., Morton, Mary A., and Anderson, Alan A. Exploratory Study of a Sequential Item Test. USAPRO Technical Research Note 129. December 1962.

No separate answer sheet is required. All scoring is automatic and cumulative. Test data are immediately available in printed form. All scoring may be automatically recorded on punch card or tape. Item scores may be differentially weighted, a feature especially desirable with branching tests to account for differences in difficulty of alternate routes to the same terminal item.

TEST PROGRAMS

Programming control is sufficiently flexible to permit use of any one of three types of test program:

1. The <u>linear</u> program provides a conventional test in which all examinees respond to the same set of test items.

2. The <u>branching</u> program provides a test in which the examinee's response determines the difficulty of the next item--a less difficult item if incorrect, a more difficult item if correct. A wide variety of difficulty patterns is possible; the initial item may be of any level of difficulty.

3. The <u>composite</u> program combines the linear and the branching programs. It provides a test in which all examinees answer a set of items, the score on which results in branching to items of an appropriate difficulty level.

Each type of program permits variation in item ordering and test length. The maximum number of items possible is limited by the capacity of the item storage device, described below. Examples of the variety of test plans possible are shown in Figures 1-5. Of especial interest is the flexibility possible with the programmed test. If it is necessary to differentiate across a wide range of ability, the initial item could be of average difficulty and the plan of Figure 1 followed. If, however, it is not necessary to differentiate among examinees below a certain level, the plan of Figure 2 would permit stopping the test if an examinee makes errors on several successively easier items. Conversely, if it is not necessary to differentiate above a certain level, a comparable plan would permit stopping the test if an examinee answers correctly several successively more difficult items. Various other plans could be adopted to provide a variety of distributions of scores and degree of differentiation.

To minimize the effects of chance selection at the beginning, two items of the same difficulty may be presented before branching occurs. If only one of the two items is answered correctly, branching is to an item of comparable difficulty; if both answers are incorrect, branching is to a less difficult item; and if both are correct, to a more difficult item. Figure 3 illustrates such a plan.

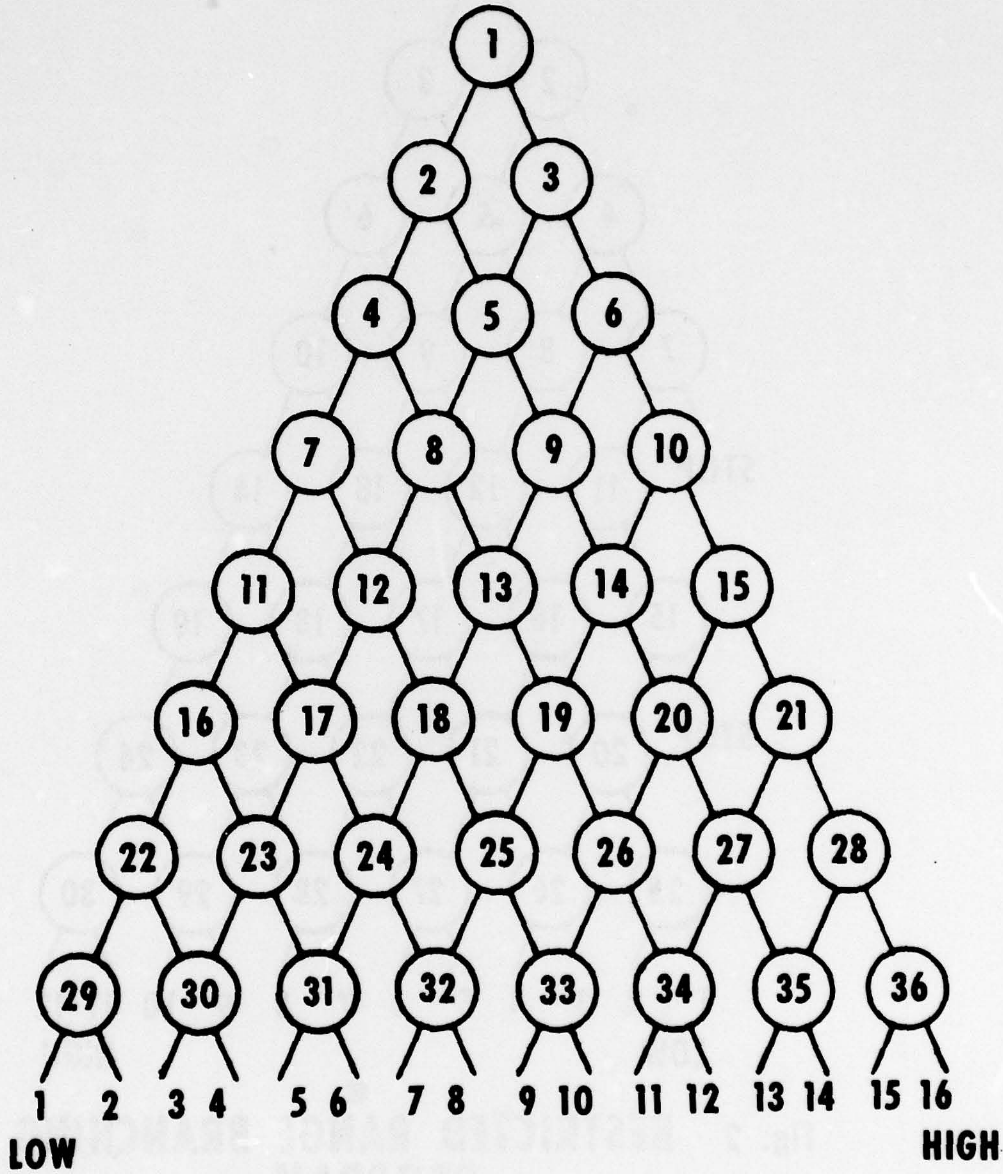


Fig. 1 **WIDE RANGE BRANCHING PROGRAM**

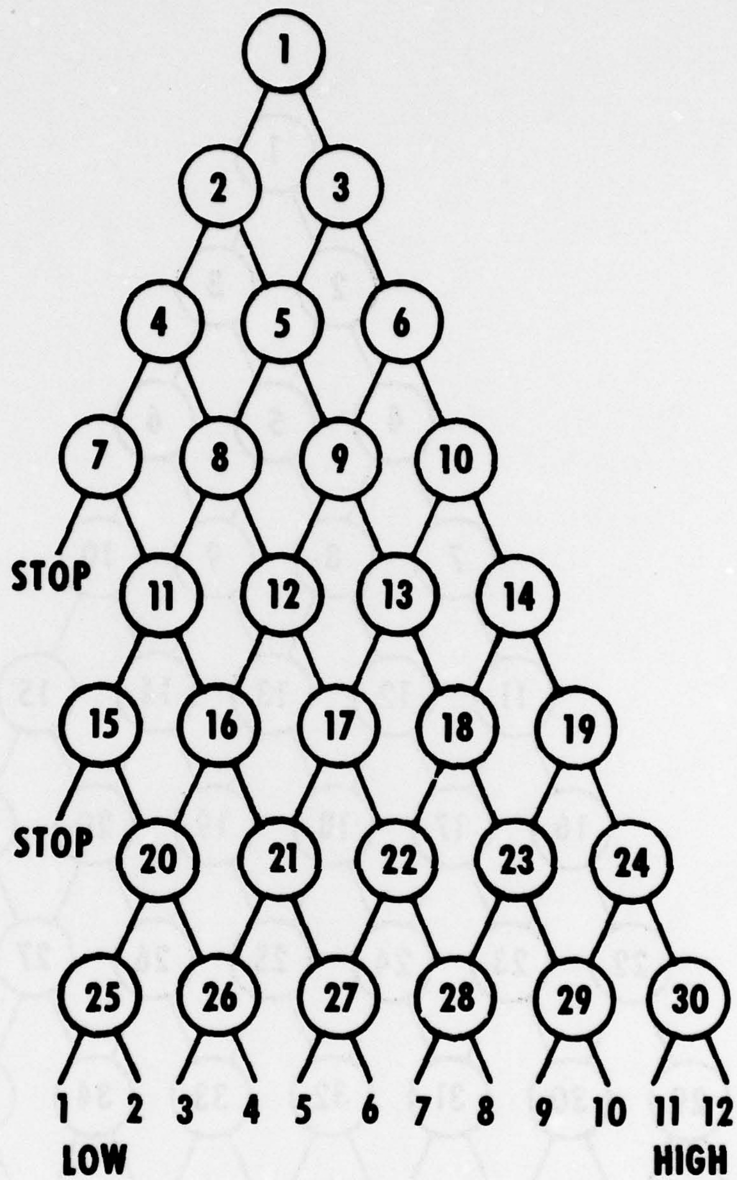


Fig. 2 **RESTRICTED RANGE BRANCHING PROGRAM**

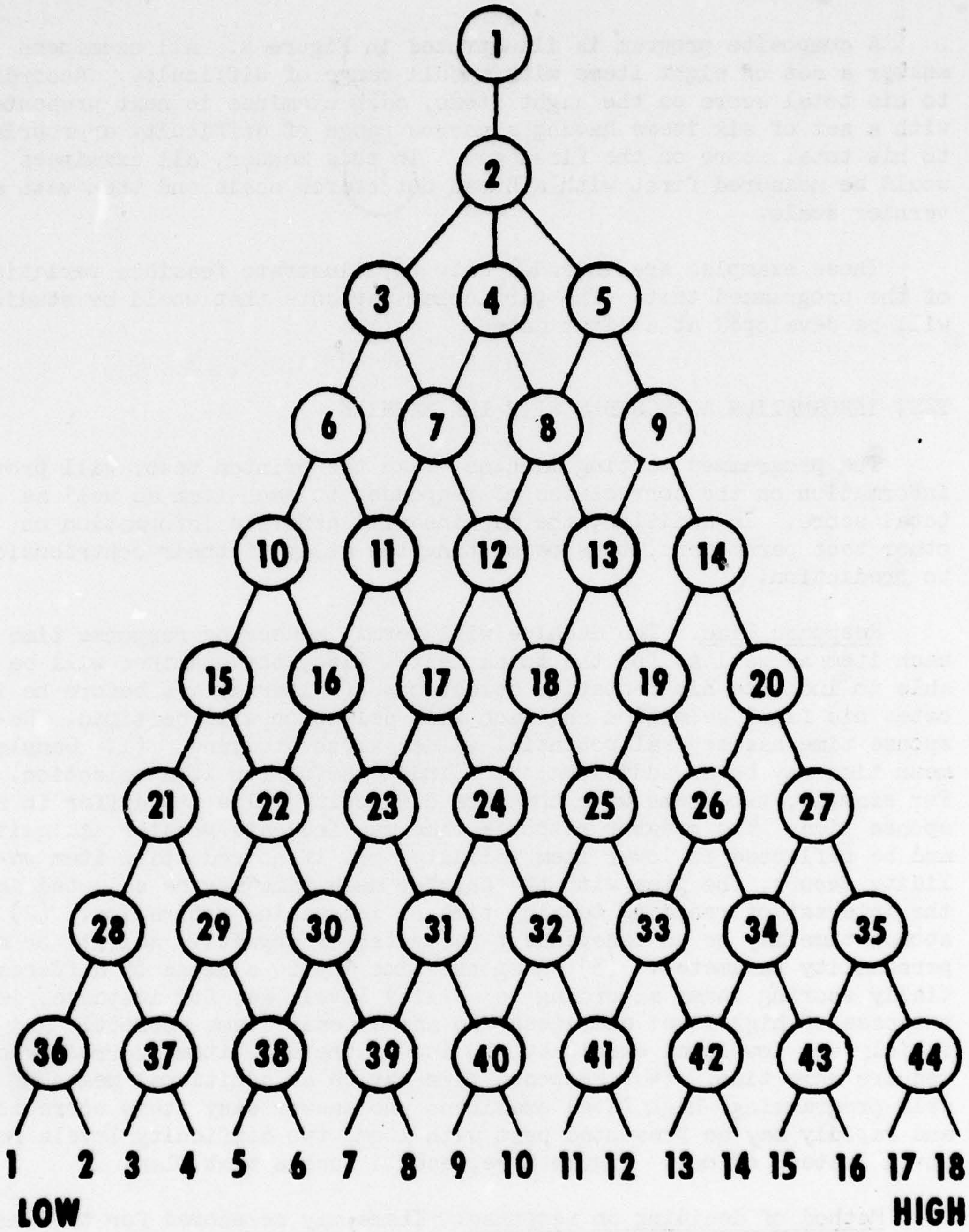


Fig. 3 **BRANCHING PROGRAM WITH REDUCED EFFECTS OF INITIAL CHANCE SELECTION**

A composite program is illustrated in Figure 4. All examinees answer a set of eight items with a full range of difficulty. According to his total score on the eight items, each examinee is next presented with a set of six items having a narrow range of difficulty appropriate to his total score on the first set. In this manner, all examinees would be measured first with a broad but coarse scale and then with a vernier scale.

These examples are intended only to illustrate feasible variations of the programmed test. The particular variants that would be studied will be developed at a later date.

TEST INFORMATION ACCESSIBLE WITH THE MACHINE

The programmed testing machine, like the printed test, will provide information on the correctness of responses to each item as well as total score. In addition, the machine will generate information on other test parameters, thus permitting the study of their contribution to prediction.

Response Time. The machine will permit measuring response time for each item as well as for the total test. Also, the examinee will be able to indicate his tentative selections of alternatives before he indicates his final selection and each such selection will be timed. Response time has several potential values worth studying. (1) Sample mean time may be an additional item index useful for item selection. For example, two items with the same difficulty value may differ in response time. The greater response time may indicate greater ambiguity and be reflected in lower item validity; or, if no reduction item validity occurs, the item with the shorter mean time may be selected in the interest of reducing testing time or increasing motivation. (2) Response time may be an independent parameter of cognitive ability or a personality parameter. (3) Response time may be a means of differentially scoring items according to ability level, as, for instance, in the case of high level examinees who answer easy items correctly and rapidly and low level examinees who answer the easy items correctly but require more time. (4) Response time may be an additional means of self-programming--high level examinees who answer easy items correctly and rapidly may be presented next with items two difficulty levels removed instead of one. Figure 5 represents such a test plan.

Method of deciding on response. Items may be scored for the number of alternatives tentatively selected by the examinee as the correct answer before the final selection is made. It would be possible, for instance, to discriminate between a correct choice made as the first choice, and a correct choice made after one or more incorrect choices were made.

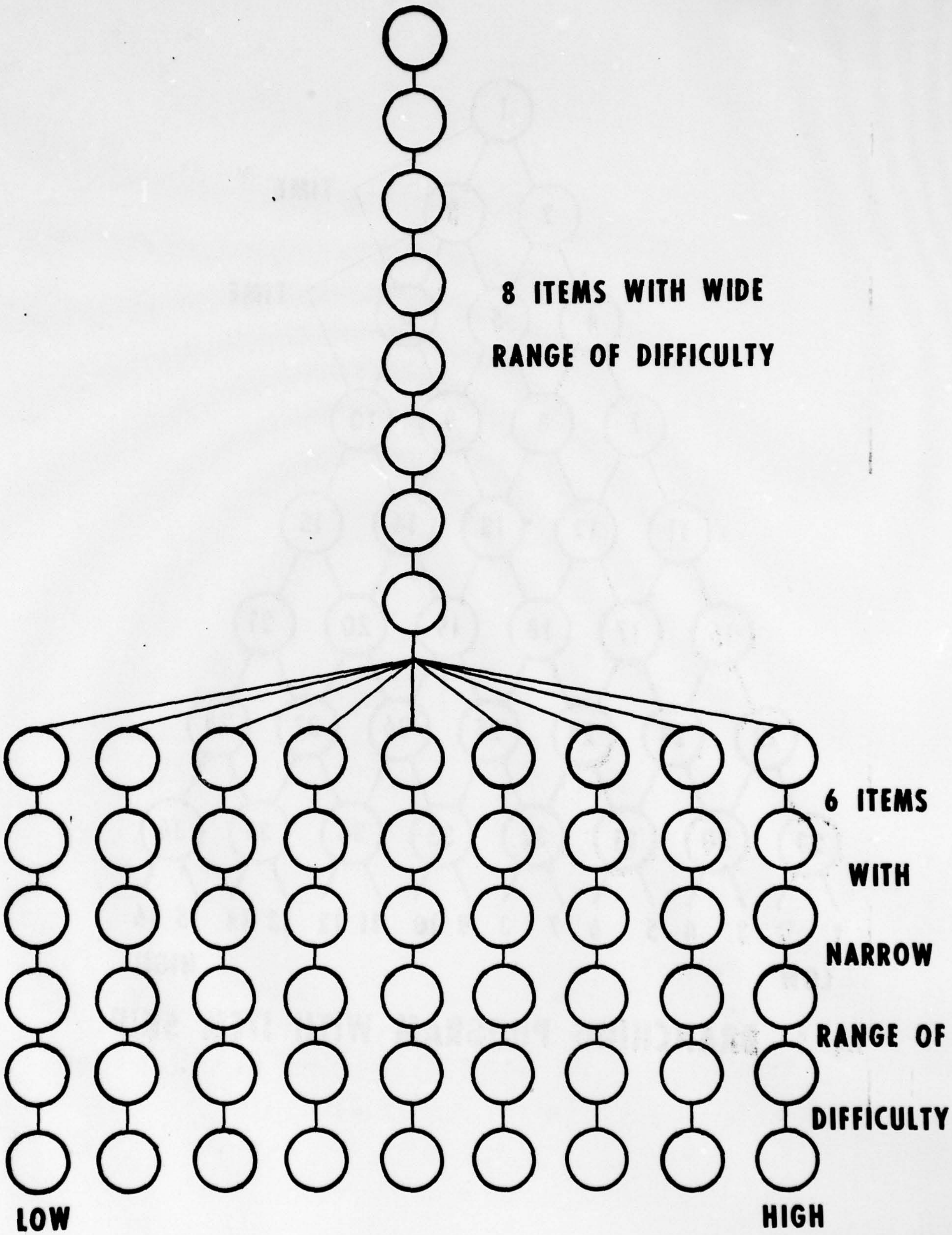


Fig. 4 **COMPOSITE PROGRAM**

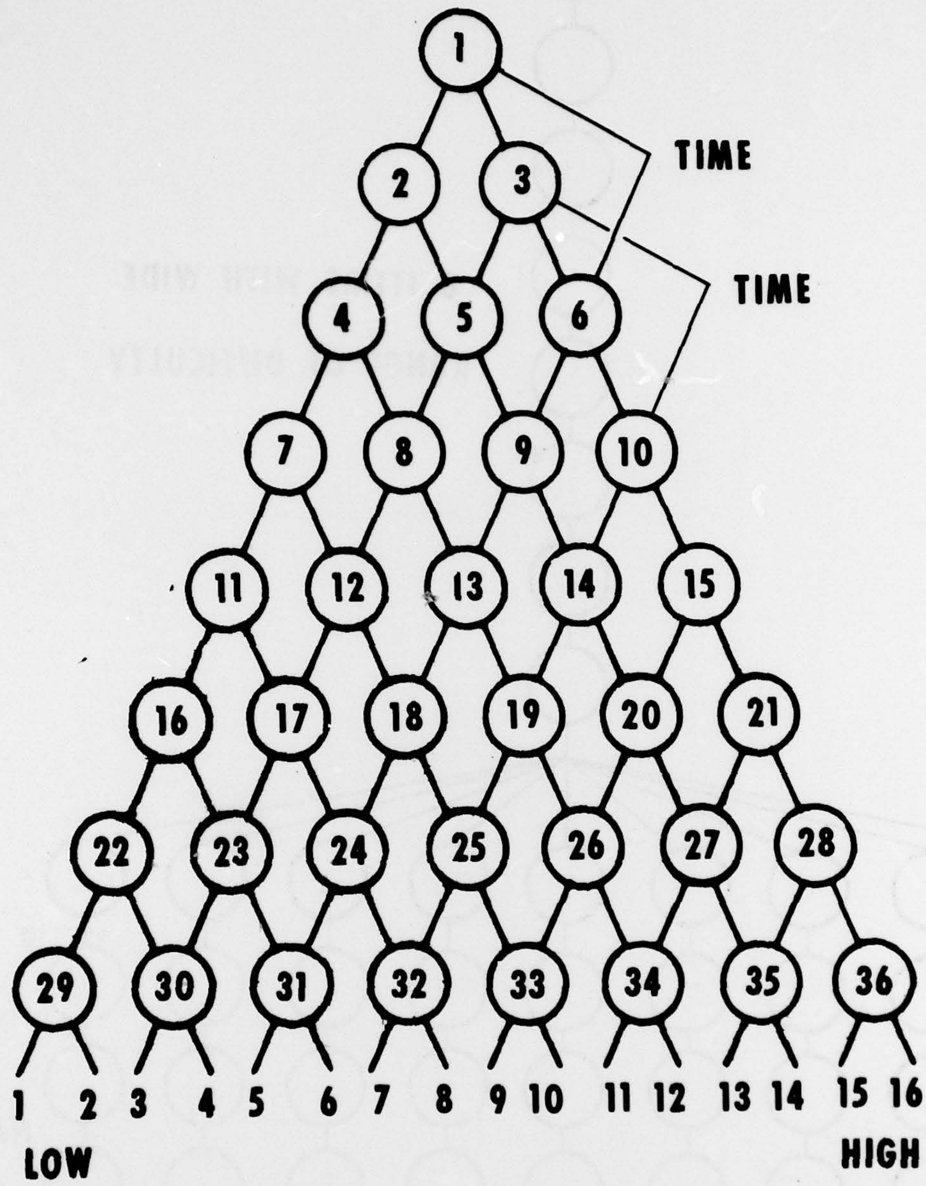


Fig. 5 **BRANCHING PROGRAM WITH ITEM SKIP**

Knowledge of results. The machine as designed may be set to provide immediate information as to the correctness of each response. In learning and psychomotor tasks, the accelerating effects of such feedback are generally pronounced. The extent to which lack of such feedback depresses scores on cognitive aptitude tests is not known. In one sense, feedback should have no effect since aptitude test items are considered to be independent of each other in content and hence knowledge of one item cannot be transferred to a succeeding item. However, it is possible that feedback may reduce invalid test variance by increasing test motivation and improving test taking methods.

The machine as designed will have the capability of providing three types of test information for the researcher, (response time, method of choosing response; for the examinee, immediate knowledge of results) for both the conventional linear test and the branching test. The contribution of this information to prediction is generally unknown. To assume that the contributions are trivial may be reasonable, but it is reasonable only because of the absence of positive evidence to the contrary. Each may be trivial but the cumulative effects may not be.

Whether or not validity of prediction is increased, the greater variety of information from conventional test items may provide the same levels of validity with fewer items. Or if there is no effect on validity, there may be a contribution to stability of norms, no insignificant problem in larger scale testing of a wide range of cultural segments of a large population over a long period of time. The aptitude test item may be a resource of information inadequately mined by conventional testing media. The testing machine is equipped to yield much that has hitherto been left buried so that it may be objectively evaluated.

OPERATIONAL TESTING

In addition to the potential value of the programmed testing machine as a laboratory tool to study methods of improving test technology, the operational version of the machine will have capabilities for testing operations and could substantially increase the efficiency of the screening system.

Design for the experimental prototype operational machine was based on a priori minimizing the capabilities of the laboratory model. The objective was to design the simplest and least expensive machine that could be used for experimentation with branching tests in an operational situation. Accordingly, the design of the laboratory model was simplified to provide the following capabilities:

1. The total pool of items to be stored would be somewhat less than in the laboratory model, with a maximum number of 10 items to be

selected for answering. The variety of test plans could be retained within these limits, with provision for study of one particular plan at a time.

2. The final score would be determined by the difficulty level of the terminal item. No differential weighting of routing would be possible.

3. Only one response per item would be permitted.

4. Response timing would be retained only if economy permits. Maximum time warning would be retained to prevent excessive delay by examinees.

This operational machine would be an interim model permitting field experimentation with branching tests. It would also permit use of conventional linear tests, since the method devised for control of item sequencing, described below, would require only that each item be appropriately coded. A final operational model could incorporate such additional features as are found useful in research with the laboratory model. If, for example, differential scoring of the route to the terminal item is found useful for prediction, this capability would be added.

Should a laboratory model be constructed, construction of a prototype operational model would entail relatively little additional expense, since most of the design and engineering costs would be absorbed in construction of the laboratory model.

MACHINE LOGIC AND COMPONENTS

A description of the machine logic and major components will provide a general understanding of methods employed for providing the capabilities described above. (See Appendix for a logic diagram and flow chart of operations.) The point of departure was the current availability of logic units and components selected so as to provide the maximum capability feasible, rather than to design new units and components to provide the maximum capability desirable.

STORAGE AND DISPLAY OF TEST ITEMS

Requirements to be met by the machine for laboratory use included means of storage and display of test items. Study of the most promising variants of the self-programming test model indicated that the laboratory machine needed a maximum pool of about 80 items out of which a smaller number would be displayed for testing. Hence, the storage capacity was set at 80 items. In addition, the following requirements had to be met:

1. Substitution of items in storage should be easy.

2. Each stored item could be displayed as readily as any other item in any position in the sequence.

3. Size of display should be adjustable within limits at reading distance.

4. Duration of display and interval between displays should be experimentally variable and, once pre-selected, constant.

5. No motion of display should be perceptible, and illumination would be adjustable for optimal visibility without elaborate shielding.

6. If possible, instructions to the examinee should be stored and displayed in the same manner as test items.

From several possible means of storage and display, a 35 mm slide projector was selected as best meeting the requirement. A commercially available projector having the necessary features was located. It contains a radial magazine with satisfactory slewing characteristics, has a capacity of 80 slides, and permits easy modification to provide the necessary controls. The magazine can be removed as a unit, or individual slides can be replaced. Each slide would contain one item.

The experimental prototype operational machine would differ from the laboratory model in that the operational machine would have a film strip projector instead of the slide projector, and a capacity of 60 items instead of 80, out of which a maximum of 10 items could be selected for display.

TEST PROGRAMMING

Item coding. In designs for both laboratory and operational models, a maximum of seven alternatives is possible for each item with provision for a reduced number if desired. Only one alternative can be keyed as correct. Any number of alternatives may be selected by the examinee, and any selection may be repeated except as limited by the program. When the examinee indicates "record", the last alternative selected is recorded as his final answer. It is also possible to permit only one choice.

Each slide in the laboratory model (frame in the operational model) would be coded for the following information:

1. A 3-bit code to indicate which of the 7 alternatives is correct.

2. A 4-bit code to indicate which one of 15 possible items is to be displayed next.

3. A 1-bit code to indicate whether responses are to be timed.

4. A 1-bit code to indicate whether branching is to occur after the response is recorded.

Physically, the code will be represented by 9 spots photographed on the edge of each slide or frame. If the spot is opaque, no light will pass through to activate a photo-cell (0); if transparent, the photo-cell will be activated (1). The spots will not be visible to the examinee.

Item storage positioning. Each item slide will be positioned in the magazine according to the test plan, and each position in the magazine will be assigned a serial number. The serial positioning of the items represents the pattern of difficulty in the item pool for the particular test plan. An item used in a variety of test plans may be placed in different positions in the magazine.

Item sequencing. In the linear test program, the sequence of items displayed is determined by the serial order of the items in the magazine. In the branching programs, the next item presented bears a higher serial number, but an addend for each item is provided to permit branching. If an incorrect response is made, the next item (less difficult) would be No. N; if a correct response, the next item (more difficult) would be No. N + 1. If the test plan requires responses to two items of the same difficulty before branching, the addend for two incorrect responses is No. N (less difficult); for one correct and one incorrect, N + 1 (same difficulty); for two correct, N + 2 (more difficult). Thus, in a branching test, the item displayed next always has a higher serial position in the magazine, but it may be several positions removed, depending on the correctness of the response and the test plan. The 4-bit code on each slide permits branching to a maximum of 15 slides in serial order. Branching to a maximum of 15 items sets a limit to the variety of test plans. However, this limit is not serious, since none of the variants of the branching test program considered for possible study exceeds this limit.

Program control. 1. Selection of the next item to be displayed is controlled by signals from the examinee's response switches, by the code read from the item slide by photo cells, and by signals from a program tape which represents the test plan (the slide addends) and item weights. The program tape is a loop of 8-channel punched tape mechanically coupled to the projector so that the tape is positioned concurrently with the slewing of the magazine to the next slide to be displayed.

2. Several time controls are provided. The maximum slewing time of the projector (i.e., from the slide in the first serial position to the slide in the last position) is 5 1/2 seconds. A control is incorporated to insure constancy of the interval between recording the final response to one item and presentation of the next item, regardless of the

distance between their positions in the magazine. The duration of display is adjustable, and once adjusted, the duration is constant.

3. A maximum time warning may be set to inform the examinee that he is nearing the time limit. If he does not record his final choice, the last choice made may be automatically recorded as the final choice. If he has made no choice, a signal may be sent to the examiner who may instruct the examinee to make some selection or allow the examinee excess time. Responses recorded at the instance of the examiner may be scored as omits.

4. An item-skip timer may be set so that an examinee who responds correctly in less than allotted time will have the next programmed item automatically recorded as correct without displaying it (i.e., the item will be skipped). The item displayed next will represent the second step in the sequence of increasing difficulty. The time interval for item skipping is variable.

DATA PROCESSING

A tally register is provided to count the number of times each alternative was tentatively chosen and to indicate the correctness or incorrectness of the recorded alternative. The examinee's activation of the "record" switch initiates the data processing cycle. During this cycle, the display and the "record" response switch are deactivated and timers stopped so that the examinee cannot disrupt the processing operations by activating any switch.

When the response recorded is correct, the correct response subtally register is increased by one; when incorrect, the register is not changed. The following data are then recorded: item number, number of times each alternative was tentatively chosen, the final alternative recorded, the correctness of the recorded alternative, and the state of the item-skip timer if activated.

The serial number of the item to be displayed next is computed by adding the addend from the program tape to the displayed slide number register. If no branching is to occur, the new number is the next serial number. If branching is to occur, the correct score subtally is added to the displayed slide number register to complete the calculation of the serial number of the slide to be displayed next. A new score is computed by adding the score (weighted) for the correct response to the cumulative score register. A second set of data is now recorded: time to final recorded response and accumulated score.

The timers and correct score subtally register are now cleared and the slide magazine is activated to display the slide bearing the serial number stored in the display slide number register. The timers and record switch are reactivated to complete the cycle.

With item skip, the timers, "record" switch, and display of next item are not reactivated after the response to the displayed item. Instead, a new cycle is automatically initiated to display the second item in the difficulty sequence and to record a correct score for the skipped item. The number of times each alternative of the skipped item was selected and the response times are recorded as zero.

When the last item has been answered and the data recorded, a slide (the slide in the last position in the magazine) is displayed indicating the end of the test.

DATA RECORDING

Data to be recorded and method of recording were determined on the basis of research needs. Certain data, particularly identification information, are to be recorded manually. In most research studies, the manual recording of identification data is a minor part of data collection. Hence, the added expense of providing for automatic recording was considered unnecessary.

The following data would be recorded automatically for each item:

1. Item number
2. Number of times each alternative was tentatively selected by examinee
3. Alternative finally selected and recorded by the examinee
4. Correctness of recorded alternative
5. Time elapsed from presentation of item to recording of final selection, including time to each tentative selection
6. The state of the item-skip timer
7. Accumulated weighted raw score

Information that would be recorded manually includes identification codes for test, examinee, date. Raw scores converted to other scales (e.g., AFQT percentiles, Army Standard Scores, etc.) would also be recorded manually, if needed.

Two recording modes may be used, print-out and punched card or tape. A 12-column printer is commercially available which would automatically print out the recorded item data on adding machine tape. The process is very rapid and a complete record would be immediately available as soon as the examinee has responded to his last item. To avoid necessity of transcribing the printed record to a form which may be fed into a computer,

a card or tape punch may be attached. The immediate availability of a punched record would permit a considerable increase in efficiency of processing of research data.

SLIDE PREPARATION AND MAGAZINE LOADING

Slides will be prepared by photographing with a microfilm camera modified to insure correct adjustment of reduction ratio and focus. The 3-bit code designating the correct answer will be photographed on the edge of the slide so that when projected, the answer code will not appear on the viewing screen. Several slides may be prepared for each item with appropriate changes in the item code so that an item may appear in several different tests without physically moving the slide from one magazine to another.

Item slides are loaded in the magazine according to a designated pattern of difficulty. Certain positions in the magazine are reserved for other information. The first position is to contain a slide with the necessary test identification. The slide in the second position is to present the instructions to the examinee and practice problems. If needed, the slides in the successive positions may continue the instructions and practice problems. The slide in the last position in the magazine will indicate end of test. When these slides are in position for projection, the program control tape will inhibit the scoring operations.

DESCRIPTION OF PHYSICAL EQUIPMENT

EXAMINEE AND OPERATOR CONSOLES

The equipment will be housed in two cabinets, an examinee's console and an operator's (examiner's) console. The examinee's console will display the slides through a rear projection screen suitably shielded from ambient light. The examinee's controls will consist of seven buttons or switches to indicate alternatives selected, one button to start display of test items after the instructions and practice problems have been displayed, and one button to record the final selection. In addition, there will be indicators, subject to inhibition by the examiner, of the last choice made and of the correctness of the recorded choice. If the test items have fewer than seven alternatives, a housing may be placed over the selection buttons to expose only the buttons corresponding to alternatives, whatever their number.

The operators's console contains several sets of controls. One set governs the preliminaries to the display of the test items. There is a control to position the magazine properly for changing magazines, another to display the test identification and instructions to the operator (set response timers, inform examinee of correctness of choice, etc.).

Several sets of dials are provided for entering numeric test and examinee identification codes and for setting the timers.

Once these controls are set and the examinee operates the "start testing" switch, the operator is free to act as proctor.

THE CONTROL AND COMPUTING SYSTEM

According to the design, the logic units are to be solid-state components in the form of printed circuit plug-in modules. Servicing of the logic units will consist of replacement of the plug-in modules. The use of solid-state components results in great reliability, quietness of operation, minimal heat dissipation, and reduced maintenance.

The only component that is not electronic is the slide selection mechanism. This mechanism will be electro-mechanical and will be driven by electrical relays.

The timers will be electronic rather than electro-mechanical. The first cost of electro-mechanical timers is less than of electronic timers. However, their general incompatibility with solid-state circuitry, awkwardness of resetting, lack of accuracy over a broad range, and need for maintenance, makes them less desirable than electronic timers.

In designing the prototype machine, emphasis was placed on reliability and simplicity of maintenance. Important as these requirements are for a laboratory tool, they are even more important in an operational machine. One of the potential contributions of a testing machine is to make the functioning of the operating screening system more efficient. Hence, little would be gained if the machine could not be depended on for sustained operation, or if scarce high-quality personnel would be needed for maintenance and servicing.

SUMMARY

A preliminary design study has indicated that a programmed testing machine can be developed which would meet research requirements and also be adaptable for operational use. The study did not include cost estimates for construction of either the research model or the operational model. However, the laboratory model could probably be procured at a cost comparable to that of other computerized research equipment.

Machine testing as a research medium can have substantial impact on psychometric experimentation. With computerized equipment, research may be directed at test techniques not accessible to study by the conventional printed test. Test data hitherto obtainable only by means of costly and time-consuming individual examination can be subjected to

validation for its contribution to prediction. Basic changes in testing technique would be opened up for study. Improved prediction of trainability and soldier performance could result. Or, as is more likely in view of the limits to the predictability of available criterion measures, the same degree of prediction could be attained with shorter tests.

Researchers on psychological measurement are reluctant to leave unexplored the potentials for improved testing instruments which reside in the application of computerized techniques. Immediate interest is therefore in the construction of the laboratory model with facilities for handling new kinds of test information and introducing tests which depart in concept from conventional printed tests. Feasibility research can then continue by empirical evaluation of the various new forms and techniques. If gains should materialize from the research in the form of improved predictors, shorter testing time, fewer administrative errors, and reduced personnel requirements for test administration, these could then be weighed against procurement costs for large-scale installation of operational machines in Armed Forces Examining Stations. In manpower terms, reduction in error of prediction would mean elimination from the training cycle of unqualified men who might be accepted with present screening measures and the acceptance of potentially useful soldiers who might otherwise have been rejected.

APPENDIX

Logic Diagram of Programmed
Testing Machine

Flow Chart of Operations

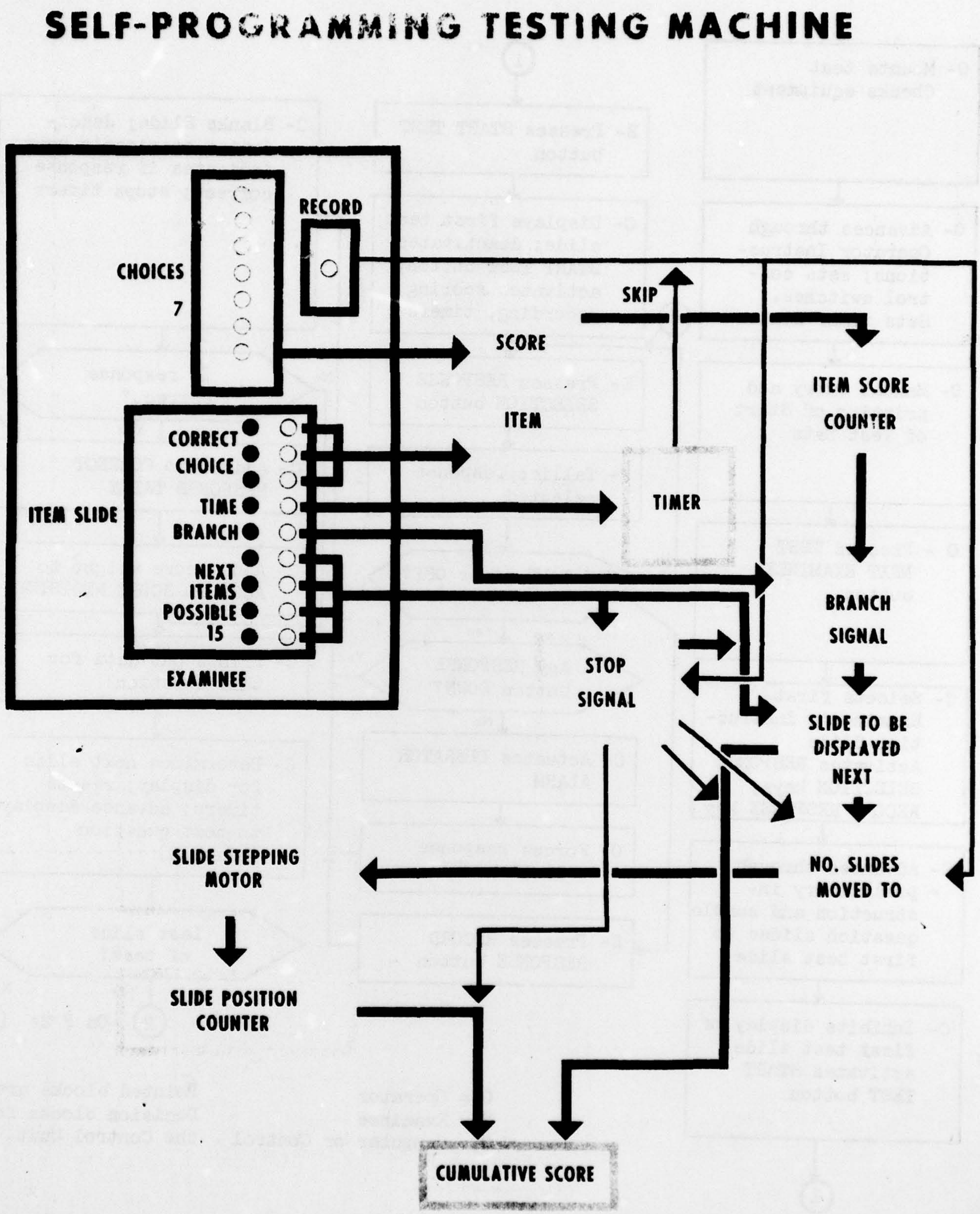


Figure A-1. Logic Diagram of Programmed Testing Machine

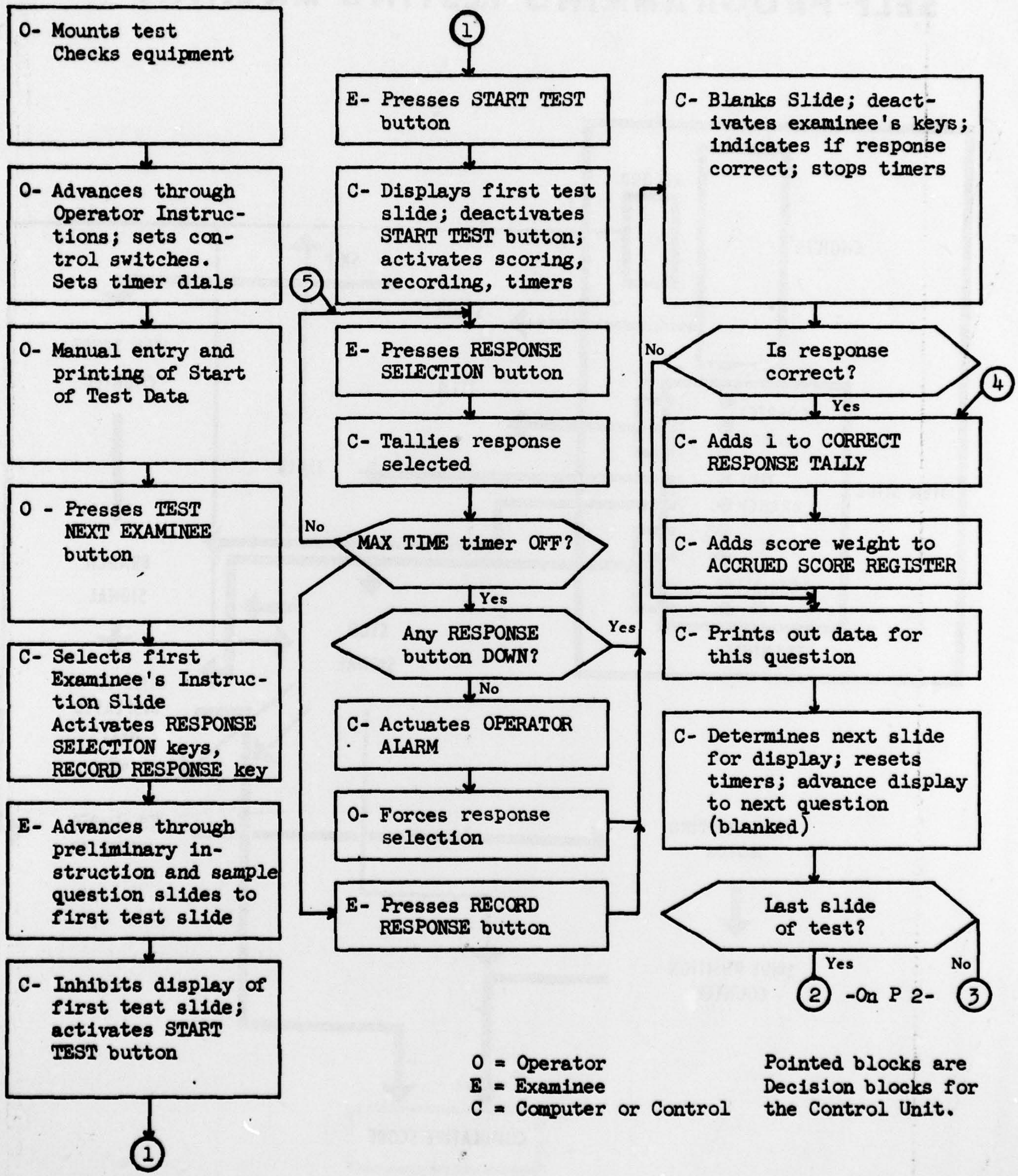


Figure A-2. Flow Chart of Operation (PAGE 1 OF 2 PAGES)

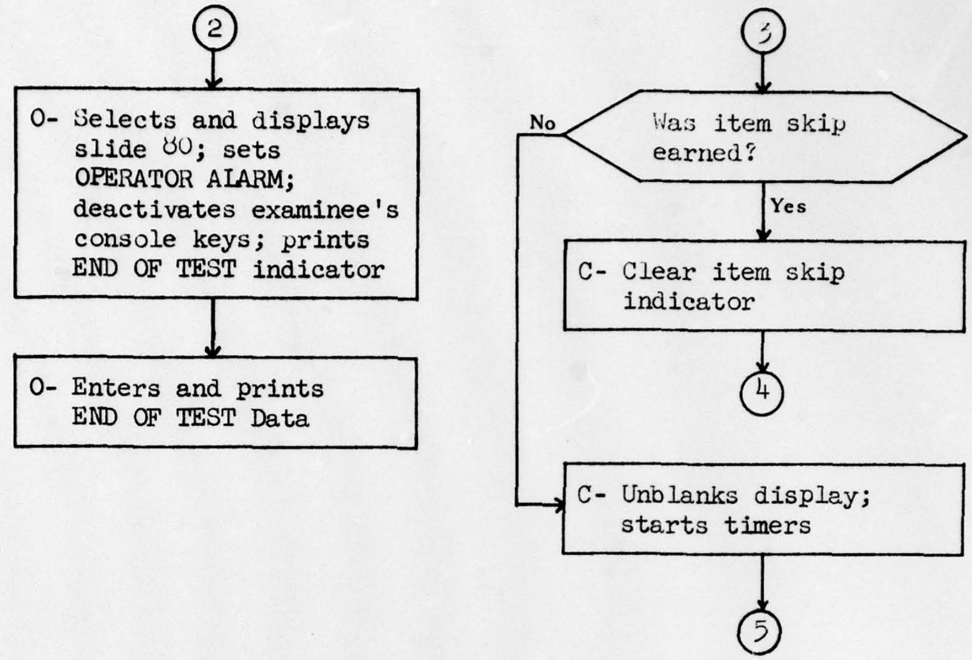


Figure A-2. Flow Chart of Operation (Cont'd) (PAGE 2 OF 2 PAGES)